

PANEL AND LONGITUDINAL DATA ANALYSIS: ASSIGNMENT 2

1 December 2024

Time : unlimited hours

Candidates should attempt ALL questions. Marks will be allocated as indicated.

A1. In a study to investigate weight loss programs, 100 male subjects were randomly assigned to maintain their current eating and exercise habits (Control), to follow a rigorous program of modified diet and exercise (Diet+Exercise), or to adopt a modified diet (Diet Alone). All participants were weighed at baseline (month 0) and then again at months 3, 6, 9, and 12. An analyst considered a model with different linear time trends for the three groups and a random intercept to account for the correlation of the measurements. The following is part of the output from R.

(a) Using symbols for fixed effects, random effects, and residual error terms, write down two expressions (i.e., one for each program) describing the trend of weight loss in the two programs. Please use clear notation, making sure that all symbols used are defined. [8]

(b) The analyst fitted the intended model in R and provided the following output

```
> summary(fitM1)
Linear mixed-effects model fit by maximum likelihood
Data: tmpDat
AIC      BIC      logLik
2595.618 2618.037 -1291.809

Random effects:
Formula: ~1 | id
(Intercept) Residual
StdDev:      18.09413 12.17425

Fixed effects: weight ~ prog + month + prog * month
Value Std.Error DF t-value p-value
(Intercept)      240.84412  3.522061 246  68.38158  0.0000
progExercise       3.35374  5.240995  60   0.63991  0.5247
month              0.54941  0.221514 246   2.48026  0.0138
progExercise:month -7.98727  0.329624 246 -24.23148  0.0000

Number of Observations: 310
Number of Groups: 62
```

- (i) Based on the output, write two explicit expressions for the estimated population average trends in each of the programs. [3]
 - (ii) Make an appropriately labelled sketch of the average trends in the two programs based on your results in b(i). Comment on your plot. [3]
 - (iii) Use the output above and any other relevant information to verify that the data collected is balanced longitudinal data. [2]
 - (iv) Provide an estimate of the correlation between any two values within a subject over time. [4]
- (c) The analyst claimed that the slopes were similar between the two groups and provided output for another model that was considered

```
> summary(fitM2)
Linear mixed-effects model fit by maximum likelihood
Data: tmpDat
AIC      BIC      logLik
2897.008 2915.691 -1443.504

Random effects:
Formula: ~1 | id
(Intercept) Residual
StdDev:      16.00934 22.44341

Fixed effects: weight ~ prog + month
Value Std.Error  DF   t-value p-value
(Intercept) 262.48704 3.726296 247  70.44182      0
progExercise -44.56987 4.845601  60  -9.19801      0
month        -3.05774 0.301914 247 -10.12785      0

Number of Observations: 310
Number of Groups: 62
```

- (i) Based on this output, write two explicit expressions for the average trends in the two groups. [5]
- (ii) Using your results in (a) or otherwise, write down a null hypothesis that addresses the analyst's claim of similar slopes in time between the two groups. [3]
- (iii) Using relevant output, perform a formal statistical test to assess the analyst's claim of similar slopes in the two groups. Comment on your results. [5]

END OF ASSIGNMENT 2 DUE 4 December 2024 16:00